Poster II-23

A Web-Based Integrated Data Management, Retrieval, and Visualization System for Assessing Climate and Environmental Impacts on Epidemic Using Remotely Sensed Data

Liang, Yao[1], Liang, Xu[2]
[1]Alexandria Research Institute, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA; [2]Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA

Remotely sensed epidemic surveillance is a new interdisciplinary research field that requires a sophisticated information technology infrastructure to efficiently retrieve, analyze, and visualize a huge amount of remotely sensed data from diverse distributed sources. Due to the wide variety of data formats and structures of existing climatological and remotely sensed data, medical scientists will have to spend a significant amount of effort to convert the data from their original heterogeneous formats and structures into usable information before they can use these data to conduct any scientific analyses. It is also difficult and very time-consuming for medical scientists to develop advanced data analyses and visualization applications by themselves alone.

In order to facilitate medical scientists to access, analyze and visualize data more efficiently and effectively, so that they can spend less time as computer programmers and therefore focus on researches in their areas, we have developed a web-based integrated distributed data management, retrieval and visualization system for medical studies. Our system covers many heterogeneous data sources (including NASA and NOAA) with different data structures and formats, and is able to manage, retrieve, share, analyze, and visualize large data volumes easily. The system is designed with full flexibility, extensibility, scalability, uniformity, transparency and heterogeneity. XML based metadata mechanism is the foundation of our data management system. The dynamically generated query GUI (Graphical User Interface) in our system makes it easy and convenient for medical scientists to access heterogeneous datasets. Scientific data visualization methods are developed to display a huge amount of data graphically to help researchers have better understanding of the data and gain valuable insights of the datasets under investigation. OPeNDAP data access protocol is employed to enable secure and reliable distributed data delivery. Without knowing any information of the physical storage location, content, structure and format of each dataset instance, and without programming a single line of codes, medical scientists can now query heterogeneous and distributed data easily, and view and understand the retrieved data in analytical and graphical ways. Further, diverse sources of medical datasets can be readily added and integrated into the current system to provide correlated analysis among remotely sensed data and epidemic datasets.